Cover Feature

Holographic Data Storage

Digital data storage using volume holograms offers high density and fast readout. Current research concentrates on system design, understanding and combating noise, and developing appropriate storage materials. Possible applications include fast data servers and high-capacity optical disks.

Demetri Psaltis
California Institute of Technology

Geoffrey W. Burr
IBM Almaden Research Center

Has the rapid increase in available storage capacity fueled the demand for storage, or vice versa? It's hard to say: Computer users' hard disk drives are perpetually overflowing with data, even though a year earlier the same-size disk seemed more than sufficient. Research into and development of data storage devices is a race to keep up with this continuing demand for more capacity, more density, and faster readout rates.

Improvements in conventional memory technologies—magnetic hard disk drives, optical disks, and semiconductor memories—have managed to keep pace with the demand for bigger, faster memories. However, strong evidence indicates that these two-dimensional surface-storage technologies are approaching fundamental limits that may be difficult to overcome, such as the wavelength of light and the thermal stability of stored bits. An alternative approach for next-generation memories is to store data in three dimensions.

In this article we discuss recent developments in holographic 3D memories,[1-4] in which researchers achieve high density by superimposing many holograms within the same volume of recording material. (For a brief look at other techniques, see the sidebar, "Methods of Optical Storage.") Data are encoded and retrieved as two-dimensional pixelated images, with each pixel representing a bit. The inherent parallelism enables fast readout rates: If a thousand holograms can be retrieved each second, with a million pixels in each, then the output data rate can reach 1 Gbit per second. For comparison, an optical digital videodisk (DVD) outputs data at about 10 Mbits per second.

The surface density that has been experimentally demonstrated in holographic memories is 100 bits per squared micron in a 1-mm-thick material.[5] This can reach in excess of 350 bits per squared micron with thicker materials.[6] By comparison, the surface density of a DVD disk is 20 bits per squared micron; of magnetic disks, 4 bits per squared micron.

The potential of this storage technology has generated a large research effort. The participants range from Bell Labs to a large consortium funded by the US Defense Advanced Research Projects Agency (which includes industrial partners such as IBM, Optitek, Rockwell, Kodak, and Polaroid), to independent start-up companies such as Holoplex.

HOLOGRAMS

A hologram is a recording of the optical interference pattern that forms at the intersection of two coherent optical beams. Typically, light from a single laser is split into two paths, the *signal path* and the *reference path*. Figure 1 shows this holographic recording arrangement. The beam that propagates along the signal path carries information, whereas the reference is designed to be simple to reproduce. A common reference beam is a *plane wave*: a light beam that propagates without converging or diverging. The two paths are overlapped on the holographic medium and the interference pattern between the two beams is recorded.

A key property of this *interferometric* recording is that when it is illuminated by a readout beam, the signal beam is reproduced. In effect, some of the light is diffracted from the readout beam to "reconstruct" a weak copy of the signal beam. If the signal beam was created by reflecting light off a 3D object, then the reconstructed hologram makes the 3D object appear behind the holographic medium. When the hologram is recorded in a thin material, the readout beam can differ from the reference beam used for recording and the scene will still appear.

Volume holograms

When a hologram is recorded in thick material, the portion of incident light diffracted into the direction of the object beam (the *diffraction efficiency*) depends on the similarity between the readout beam and the original reference beam. A small difference in either the wavelength or angle of the readout beam is sufficient to make the hologram effectively disappear. The sensitivity of the reconstruction process to these small variations in the beam increases, approximately linearly, with material thickness. Therefore, by using thick recording

0018-9162/98/$10.00 © 1998 IEEE

media, designers can exploit this angular or wavelength readout sensitivity to record multiple holograms.

To record a second, angularly multiplexed hologram, for instance, the angle of the reference beam is changed sufficiently so that the reconstruction of the first hologram effectively disappears. The new incidence angle is used to record a second hologram with a new object beam. The two holograms can be independently accessed by changing the readout laser beam angle back and forth. For a 2-cm hologram thickness, the angular sensitivity is only 0.0015 degrees. Therefore, it becomes possible to store thousands of holograms within the allowable range of reference arm angles (typically 20–30 degrees). The maximum number of holograms stored at a single location to date[7] is 10,000.

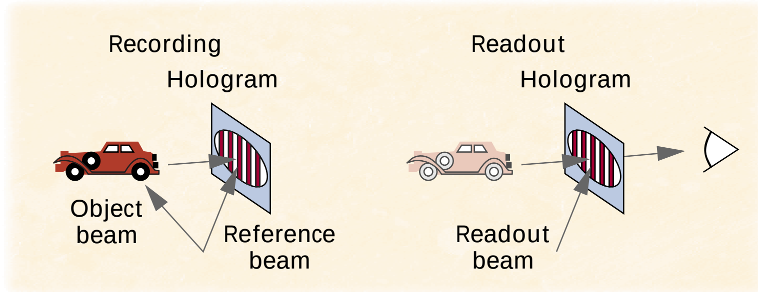

Figure 1. After recording a hologram with object and reference beams, a readout beam can be used to reconstruct the object beam and make the object reappear to a viewer.

Storing and retrieving digital data

To use volume holography as a storage technology, the digital data to be stored must be imprinted onto the object beam for recording, then retrieved from the reconstructed object beam during readout. The input device for the system is called a *spatial light modulator*, or SLM. The SLM is a planar array of thousands of pixels; each pixel is an independent optical switch that can be set to either block or pass light. The output device is a similar array of detector pixels, such as a charge-coupled device (CCD) camera or CMOS pixel array. The object beam often passes through a set of lenses that image the SLM pixel array onto the output pixel array, as Figure 2 shows.

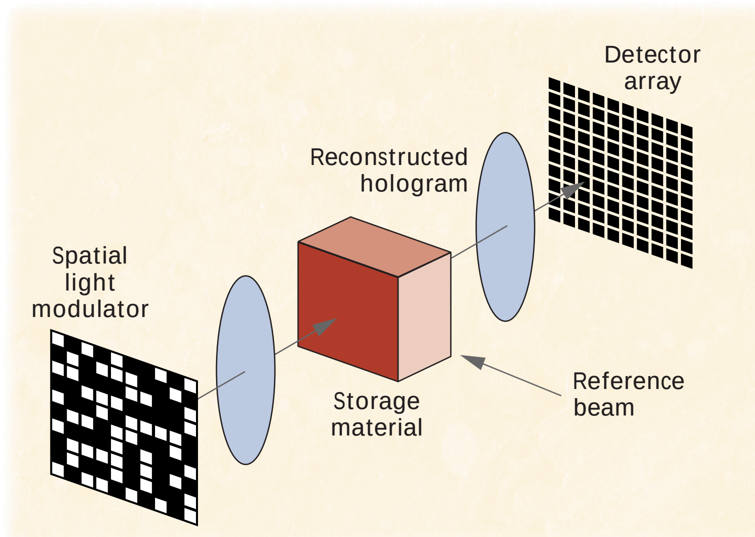

Figure 2. For data storage, information is put onto the object beam with a spatial light modulator and removed from a reconstructed object beam with a detector array.

Methods of Optical Storage

Optical data storage techniques are categorized in three basic groups. The symbol • that precedes a technique indicates a method that is already in use to produce commercial products.

Surface or 2D recording

• **CD/DVD**—Data are stored in reflective pits and scanned with a focused laser. Disks are easily replicated from a master.

• **CD-Recordable**—Reflective pits are thermally recorded by focused laser. This type is usually lower density than read-only versions. Researchers have proposed blue lasers and "electron-trapping" materials to achieve density improvements.

• **Magneto-optic disks**—Spots are recorded with a combination of magnetic field and focused laser.

Near-field optical recording—Higher 2D density than with conventional surface recording is achieved by placing a small light source close to the disk. Light throughput and readout speed are issues.

Optical tape—Parallel optical I/O has the advantages of magnetic tape without the long-term interaction between tape layers wound on the spool. Flexible photosensitive media is an issue.

Volumetric recording

Holographic—Data are stored in interference fringes with massively parallel I/O. Suitable recording material is still needed.

Spectral hole burning—This technique addresses a small subset of molecules throughout the media by using a tunable narrowband laser. Alternatively, all subsets are addressed with ultrashort laser pulses. It may add a fourth storage dimension to holography but requires cryogenic temperatures and materials development.

Bit-by-bit 3D recording

• **Sparsely layered disks**—The focus of the CD laser is changed to hit interior layers. DVD standard already includes two layers per side.

Densely layered disks—A tightly focused beam is used to write small marks in a continuous or layered material; read with confocal (depth-ranging) microscope.

2-photon—Two beams of different wavelengths mark writes, then read in parallel using fluorescence. Material sensitivity is an issue.

Most holographic read-write materials are inorganic photorefractive crystals doped with transition metals such as iron or rare-earth ions such as praseodymium, grown in large cylinders in the same way as semiconductor materials.

The hologram can be formed anywhere in the imaging path between the input pixel array and the output pixel array. To maximize storage density, the hologram is usually recorded where the object beam attains a tight focus. When the reference beam reconstructs the hologram, the object beam continues along the original imaging path to the camera, where the optical output can be detected in parallel and converted to digital data. Capacity and readout rate are maximized when each detector pixel is matched to a single pixel on the SLM, but for large pixel arrays this requires careful optical design and alignment. Recently, the IBM Almaden Research Center built two test platforms that implement this pixel-to-pixel matching for input data masks of 1,024 × 1,024 pixels,[8] as well as for smaller real-time SLMs.[9]

STORAGE MATERIALS

Photosensitive materials for volume holography are generally classified as either read-write or write-once.

Read-write materials

Most holographic read-write materials are inorganic photorefractive crystals doped with transition metals such as iron or rare-earth ions such as praseodymium, grown in large cylinders in the same way as semiconductor materials. Large samples can be cut and polished, making thick holograms possible. These materials react to the light and dark regions of an interference pattern by transporting and trapping photo-ionized electrons.

Through the linear electro-optic effect exhibited by these crystals, the electrical fields created by the trapped charge give rise to an *index* or *phase grating* suitable for diffracting light. Thus, the spatial variations in light intensity present in the interference pattern become identical variations in the index of refraction. The trapped charge can be rearranged by subsequent illumination, which makes it possible to erase recorded holograms and replace them with new ones. However, the ease of charge re-excitation also results in the gradual erasure of stored holograms during normal readout. In the dark, the lifetime of these holograms ranges from months to years as the trapped charge slowly leaks away.

Recorded holograms can be "fixed" (made semipermanent and resistant to erasure during readout) through thermal or electronic processes. The fixing process affects all the stored holograms within a volume simultaneously. Thus, individual pages of data cannot be erased and replaced this way.

An alternative for achieving nonvolatile storage in photorefractive materials is to record at a light wavelength not normally absorbed by the crystal except in the presence of a third "gating" beam of different wavelength. This beam is present only during the recording and is switched off for readout.

Organic photorefractive polymers have also been developed. These materials provide more opportunity for performance tuning because you can fabricate them using a wide variety of constituents. However, these materials tend to be limited in thickness and require large applied voltages.[4]

Write-once materials

Writing permanent volume holograms generally involves irreversible photochemical reactions, triggered by the bright regions of the optical interference pattern. For example, a photopolymer material will *polymerize* (bind short monomer chains together to form long molecular chains) in response to optical illumination. In contrast, the molecules in a photochromic material undergo a change in their absorption behavior. Such materials are inexpensive to make in quantity. However, both types can have problems reproducing the object beam faithfully—the photopolymer because of shrinkage, the photochromic because of oversensitivity to average local intensity.

Careful system design can minimize these problems. One advantage of a photopolymer is that after recording, any leftover monomers can be disposed of without affecting the recorded holograms. A photochromic material, however, requires a separate chemical or optical step to disable the unused absorbing molecules after the holograms are recorded.

Currently available versions of these write-once materials are thin (approximately 100 μm)—the difficulties in making thick samples include insufficient optical quality or excessive absorption. As we will show later, however, new multiplexing techniques for thin materials have made write-once photopolymers one of the leading candidates for the first holographic memory products.

Dynamic range

In the readout process, the reconstructed hologram is imaged onto the output detector array, where the digital data is extracted from the detected signal. Noise can cause errors to occur during the detection process in various ways.

The basic trade-off in volume holography is caused by the fixed noise floor and the finite dynamic range of the recording material. In other words,

- the electronic detection process at the camera contributes the same amount of noise, no matter how bright the hologram, and
- as the number of holograms or the readout rate

increases, the amount of power diffracted into each hologram reconstruction decreases.

Even if all other noise sources are negligible, there will be a certain hologram strength at which the signal-to-noise ratio is inadequate for error-free detection. The number of detected electrons per pixel can be written as

$$\eta_{\text{electrons}} \;\alpha\; M/\#^2 P_{\text{readout}} \frac{t_{\text{readout}}}{M^2 N_{\text{pixels}}}, \qquad \textbf{(1)}$$

where M is the number of multiplexed holograms, N_{pixels} the number of pixels per hologram, t_{readout} the integration time of the camera, P_{readout} the power in the readout beam, and $M/\#$ a material/system constant, which measures dynamic range.[10] The storage capacity is MN_{pixel} and the readout rate is $N_{\text{pixel}}/t_{\text{readout}}$. An increase in either of these parameters leads to a decrease in the number of signal electrons.[10] Given the minimum acceptable number of signal electrons per pixel, we can maximize the capacity and readout rate by increasing P_{readout} or $M/\#$.

Different processes determine the $M/\#$ constant in photorefractives and write-once media. In a photorefractive crystal, the holograms' recording exposures must be carefully scheduled to record equal-strength holograms.[10,11] The first hologram is made quite strong. This first hologram erases slowly while the other holograms are stored, and finishes at the same strength as the weakly written final hologram. Alternatively, all the holograms can be cycled several times.[4] The equalized diffraction efficiency falls as one over the square of the number of holograms, with the $M/\#$ as the proportionality coefficient:

$$n = \left(\frac{M/\#}{M}\right)^2 \qquad \textbf{(2)}$$

The $M/\#$ constant in a photorefractive material becomes large if the holograms can be recorded faster than they erase.[7,11] In iron-doped lithium niobate, a typical $M/\#$ might be 1. This implies that to store 1,000 holograms with 1 million pixels and read each in 1 millisecond, we need about 1W in the reference beam.

A write-once material has much in common with photographic film: After a finite amount of input energy, the material is completely exposed. Each hologram gets its share of the dynamic range as it is recorded, preserving the bright and dark regions of the interference fringes. For instance, in a photopolymer material, the photosensitivity saturates as the available supply of monomers is exhausted. It turns out that the diffraction efficiency of individual holograms, when M of them are multiplexed in a saturable medium such as a photopolymer, also follows the $(M/\#/M)^2$ relationship. The most commonly used polymer is DuPont's HRF-150. The 100-micron-thick version has a $M/\#$ of 6.5,[5] which reduces the required readout power by a factor of 40.

A write-once material has much in common with photographic film: After a finite amount of input energy, the material is completely exposed.

Understanding noise

In addition to detector noise, other factors can cause errors:

- *The readout conditions change.* This can occur, for instance, when the recording alters the recording material properties. This causes unwanted changes in the reference beam path between the time the hologram is recorded and the time it is reconstructed. Often, the reference beam angle or wavelength can be tuned to optimize the diffraction efficiency and partially compensate for this effect.
- *The detector array doesn't line up with the pixel array in the reconstructed hologram.* This includes errors in camera registration, rotation, focus, tilt, and image magnification.
- *The detector is receiving undesired light*, either from light scattering off the storage material, crosstalk from other stored holograms, or crosstalk between neighboring pixels of the same hologram. Although crosstalk contributions scale with the strength of the holograms, scattering depends only on readout power and the components' optical quality.
- *There are brightness variations across the detected image.* This can be a problem if a single threshold is used across the image to separate the pixels into bright and dark and to assign binary values. These fluctuations can be caused by the SLM, the optical imaging, or the original laser beams.

COMBATING ERRORS

Commercial storage products have user error rates as low as 10^{-15}. For a 1-Gbyte hard drive, a user might expect to read the entire drive 100,000 times before a single bit error. On the other hand, at the hardware level, errors occur at a rate of perhaps 1 in 10,000 bits read. The designers decrease the error rate that the user sees by storing redundant bits along with user data. The sequence of user and redundant bits forms an *error correction code*. ECC algorithms performed in hardware after the read head can detect if a few bits within each code word are in error and then pass on corrected user bits.

The redundant bits cause a slight sacrifice in the capacity of each individual hologram. This sacrifice is measured by the ECC code rate—the fraction of bits stored that are actual user data. However, the increase in the raw *bit error rate* (BER) that can be tolerated allows additional pages to be stored, which increases overall capacity.[12] A typical ECC code with a code rate of 0.9 can handle an input data stream with a raw BER of 10^{-4}. It can output the user's data, stripped of redundant bits, with the desired BER of 10^{-15}. Since the ECC code rate drops quickly as the expected raw BER climbs above 10^{-3}, most holographic memory designs aim for a raw BER of 10^{-4} or so.

Keeping the raw BER at this target value takes careful engineering of the optical system combined with signal processing and modulation coding. Careful engineering of the optics alone is not generally the most cost-effective solution to reducing the BER to the target value.

Signal processing

Signal processing works by considering the storage device as an imperfect transmission "channel" for data that tends to smear together the signal energy from multiple bits of user data. Knowledge of how this intermixing occurs can be applied at the output end to eliminate the crosstalk and reproduce the originally transmitted bit sequence. In a telecommunications application or a bit-serial storage device like a hard drive or DVD disk, the smearing takes place between signals adjacent in time. In holographic storage, the smearing occurs spatially in two dimensions, as light intended for a particular CCD pixel diffracts into neighboring pixels.

Signal processing techniques for holographic storage are therefore 2D extensions of the 1D techniques developed for bit-serial devices. Examples of signal processing techniques used in holographic memories[4] are adaptive thresholding and normalization, equalization, filtering, and partial response precoding at the input.[13]

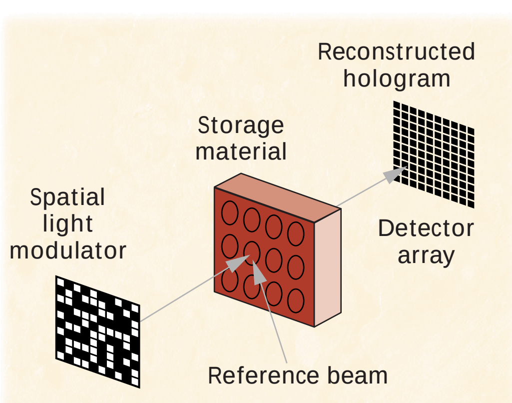

Figure 3. An HRAM system increases capacity by steering the object and reference beams to one of many storage locations on a large block of holographic recording material.

Modulation codes

A modulation code dictates the way in which bits of information are encoded into the channel as data signals. They are selected to facilitate the detection process and hence improve overall performance. For instance, in bit-serial devices, modulation codes are used to set upper and lower bounds on the frequency at which the signal level changes. In holographic storage, modulation codes are used to avoid pixel combinations that are prone to distortion and to create easy-to-detect pixel patterns.

A convenient encoding that facilitates detection is the organization of binary data into small blocks of pixels, such that the number of bright pixels is constant (usually half the pixels). The simplest example is differential encoding, in which 2 pixels convey 1 bit of information. This technique was used for storing digital data by the group at Stanford University.[3] Several modulation codes with higher code rate and performance, and thus higher complexity, have since been developed for holographic storage. These codes have been used to demonstrate as many as 1,200 superimposed holograms in lithium niobate ($LiNbO_3$) at a raw BER of 10^{-8}.[9]

SYSTEM CONFIGURATIONS

The storage capacity for the simple, angle-multiplexed memory we have been discussing is simply the product of the number of holograms superimposed times the number of pixels in each page. The number of pixels per page is currently limited to roughly one million. The dynamic range (*M/#*) of available materials and the desire for a reasonable readout rate limit the number of holograms to about 1,000. Therefore, the capacity of a single angle-multiplexed holographic memory module is only 1 Gbit.

In most cases, we must increase the capacity to at least 1 terabit to have a system that is competitive with alternative technologies. We can accomplish this by constructing a large memory consisting of multiple 1-Gbit modules. This technique is called spatial multiplexing, because multiple "stacks" of holograms are stored in different spatial locations of the recording material. Spatial multiplexing has several configuration options: holographic random-access memory (HRAM), compact modular holographic memory, and holographic 3D disks.

Holographic random-access memory

One approach for spatial multiplexing steers the reference and object beams to a stationary block of material containing multiple storage locations, as Figure 3 shows.[1,7] With nonmechanical optical scanners, the HRAM system can very rapidly steer the optical beams. Most nonmechanical beam steerers use either an acousto-optic deflector or a one-dimensional liquid crystal SLM. By using large lenses (not shown in Figure 3), the information stored at separate locations can be directed back to a single detector array.

An HRAM system can read out holograms from any location in an essentially random sequence. To maximize the number of holograms in each location, designers generally envision HRAM systems with thick read-write materials such as photorefractive crystals. Researchers at Caltech built a 16-location HRAM system capable of 10,000 holograms per location[1,7]; researchers at Rockwell demonstrated an HRAM system with no moving parts.[2]

To construct a Tbit memory using this approach, we need 1,000 spatial locations (arranged in 2D as a 33×33 array), with each location storing 1 Gbit. The main challenge in building such a system is the optics that have to simultaneously transfer data from each of the 1,000 recording sites on the recording material to a single detector array. This will require considerable engineering improvements over present systems.[2,7]

Recording rate. The photosensitivity of most photorefractive crystals is relatively low; therefore, the recording rate is invariably one to two orders of magnitude slower than the readout rate. In addition, it is practically impossible to change the state of a single pixel within a stored hologram, and it is possible but not easy to replace a single hologram within a hologram stack.[4] Instead, an entire stack of holograms must be erased together, either by heating or by illumination with the "gating" light. Therefore, an HRAM system is not truly a read-write memory; more accurately, it is an erasable write-once, read-many memory.

Readout rate. The readout rate in an HRAM system is mostly limited by the camera integration time. The reference beam reconstructs the same hologram until a sufficient number of photons accumulate to differentiate bright and dark pixels. An oft-mentioned goal is an integration time of about 1 millisecond, leading to 1,000 data pages per second. If there are 1 million pixels per data page, then the readout rate is 1 gigapixel per second. In the HRAM system this data rate can be supported continuously, independent of readout order. The latency is dominated by the integration time and is typically about 1 millisecond. This can be reduced by using a pulsed laser or a CMOS detector array or both.

Applications. The HRAM system is best matched to applications with high capacity and fast readout rate demands, yet with relatively infrequent changes to the stored data. Video-on-demand and Web servers fit well here: movie and Web content change infrequently, yet multiple users are continuously accessing enormous amounts of content in a fairly random order. (Playing one movie for a single user is sequential—playing 10 movies at once is not.)

An alternative method for reaching 1 Tbit in storage capacity is a jukebox-type apparatus. Blocks of material, each containing 1 Gbit or more, are brought into position in front of the reference beam optics for readout. The access time to a hologram is either 1 millisecond if the hologram is in the current material block, or several seconds if it is in a separate block. This can be reduced somewhat by having several readout stations. A principal advantage of increasing the capacity in this way is in the cost per megabyte of storage. For a one-block HRAM system, the cost is dominated by the components: camera, SLM, laser, beam steerers, and optics. The advantages provided by the lack of moving parts are probably enough to support this cost per megabyte only for military applications. For the commercial market, however, the cost per megabyte drops rapidly as more blocks are used, until the cost of the material becomes dominant.

The capacity of a single angle-multiplexed holographic memory module is only 1 Gbit. In most cases, we must increase the capacity to at least 1 Tbit to have a system that is competitive with alternative technologies.

Associative retrieval. An HRAM system also lets designers use a unique feature of holographic storage: associative retrieval. To search a conventional storage device for all data records sharing a particular feature, we would retrieve each record into RAM, search it using software, and continue until all records were recalled and checked. With holographic storage, this process can be performed at the memory itself.

Instead of reconstructing signal data pages with a reference beam, the data pattern of interest is put on the SLM and illuminates the storage location with the signal beam.[4] All the reference beams used to store holograms in that stack are reconstructed. The brightness of each beam, however, is proportional to the correlation between the original stored data pattern and the data pattern from the interrogating signal beam. Once the reference beams are focused onto a "correlation plane" detector array, the reference beam angles corresponding to the closest matches can be identified. The reference beam can then be used to reconstruct the desired data page onto the output camera, completing the search-and-retrieval process in perhaps 5 milliseconds.

Compact modular holographic memory

One drawback in the HRAM system is that the number of rapidly accessible locations (and thus the immediately accessible capacity) is limited by the beam-steering optics. Rather than bring the beams to the storage material, another approach is to bring the pixel arrays for data input and output to the storage material. In fact, by applying a unique feature of the stored holograms, the same pixel array can be used for both input and output. Upon readout, instead of bringing back the same reference beam used during

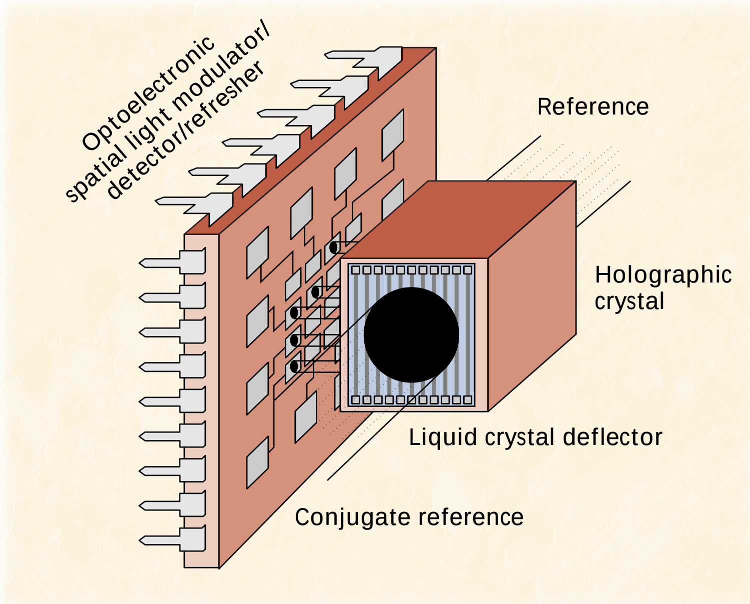

Figure 4. A compact modular holographic system can be formed by placing the holographic material close to a pixel array that performs both input and output.

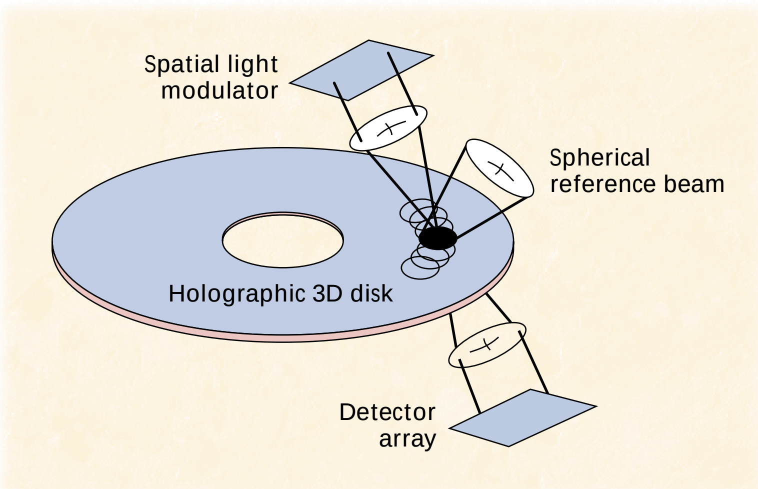

Figure 5. A shift-multiplexed disk resembles a compact disk, with an SLM and detector array for parallel input/output, and a simple reference beam. Track position and disk rotation enable access to densely overlapped holograms.

recording, its "phase conjugate" is directed to the storage location. (The phase conjugate of an optical beam passes backward along the beam's path, like a movie played in reverse.) This new readout beam reconstructs the phase conjugate of the signal beam, which returns along the original signal path back to the SLM. Because of this, a phase conjugate signal beam allows the use of a cheap imaging lens, or even no lens at all. If each pixel of the SLM is not only a light modulator but also a detector, then the entire storage device can be fabricated from identical compact modules with no moving parts. One such module is shown in Figure 4. With several modules, the memory resembles a board of DRAM, with holography increasing the amount of data stored per RAM chip.

Caltech researchers recently demonstrated a single-module compact holographic system with 480 modulator/detector pixels and 25 stored holograms.[14] A small amount of logic at each "smart" pixel let the system counteract the erasure in a photorefractive crystal by periodically detecting and refreshing the holograms. This can remove the need to fix the stored holograms and makes it feasible to individually erase holograms from a stack.

In a modular system, the cost per megabyte is dominated by the smart pixel array and the two compact angle steerers (one for the writing beam, one for the readout). Associative retrieval is possible, but adds another detector array per module. Increasing the number of pixels while keeping the cost of the detector array and angle tuners low is the key to practical implementation of this architecture.

Holographic 3D disks

A third approach to spatial multiplexing leaves the optics and components stationary and moves the storage material. The simplest method to do this employs a disk configuration.[1] The disk is constructed with a thick layer (approximately 1 mm) of the holographic material. Multiple holograms can be stored at each location on the disk surface. These locations are arranged along radial tracks, with the motion of the head selecting a track, and the disk rotation providing access along each track. As the medium thickness increases, the number of holograms that can be stored increases (and therefore the surface density goes up). However, the surface area that is illuminated also increases (and therefore the surface density goes down). Typically, it would be desirable to fabricate disks of 1-mm thickness, yielding a density of approximately 100 bits per squared micron.[6] Even though the data is still stored in 3D, in the disk configuration the surface density, not the volume density, is what matters for most practical purposes.

Angle multiplexing can be used to multiplex holograms on the holographic disk in a manner similar to the HRAM architecture. However, the angle scanner would make the readout head too large and heavy for rapid access to holograms on different radial tracks. A single, simple reference beam that could attain the same density without a bulky beam deflector would be more convenient. This can be done by making the reference beam a spherical or converging beam. This is effectively equivalent to bringing in all the reference beam angles simultaneously instead of one at a time. In this case the reconstruction becomes very sensitive to the recorded hologram's position instead of the reference's angle of incidence. In fact, if the material is shifted by a few microns, the reconstruction disappears and a second hologram can be stored. The motion of the material relative to the illuminating spherical beam needed to reconstruct different shift-multiplexed holograms is conveniently supplied by the disk's rotation. In addition, the simplicity of the reference beam makes the readout system look like a CD/DVD disk, albeit with parallel readout, as shown in Figure 5. The shift-multiplexing technique also depends on material thickness. Experimentally, it has been demonstrated that a thickness of 1 mm is sufficient to support 100 bits/μm^2, 20 times the areal density of a single-layer DVD.[5]

Holographic disks can be configured as either a WORM or a ROM system. A WORM system incorporates an SLM, turning the read head into a read-write head. The recording procedure is complicated by the chemical reactions in the photopolymers that are the recording material for the 3D disk. These reactions are not driven by light so much as triggered by it. Once begun (within an illuminated region), the reaction continues after the optical exposure stops. Because each hologram position in a shift-multiplexed WORM disk overlaps many others, the entire disk has to be recorded without stopping in order to reach the maximum capacity.

ROM applications, where the user buys a written disk (movies, audio, a computer game) and owns a simple read-only unit, may be best suited for the shift-multiplexed disk. The ROM system requires a master disk that is very similar to the WORM disk. Since the mastering device is not sold to users, it can be a large, expensive apparatus. The master disk is copied by bringing a blank disk in contact with the master and illuminating the two disks together.

Holographic disks can be viewed as a candidate technology to succeed the recently introduced DVD. They offer both higher storage density and speed. The primary factor preventing commercialization of holographic disks is the lack of a photopolymer material with sufficient thickness. Currently available materials have a thickness of 100 microns, yielding a surface density of only 10–20 bits per squared micron, which is the same order as the DVD. Research in material development promises to increase the thickness to 1 mm, increasing the achievable storage density by an order of magnitude.

Holographic storage is a promising candidate for next-generation storage. Recent research has demonstrated that holographic storage systems with desirable properties can be engineered. The next step is to build these systems at costs competitive with those of existing technologies and to optimize the storage media. If suitable recording materials become available from the research efforts currently under way, we envision a significant role for holographic storage. Three specific configurations illustrate possible examples of future holographic memories:

- *Erasable write-once, read-many drives* supporting terabytes of storage, 1 Gbit/second readout rates, and fast access to data in blocks of 50–100 Gbytes. Suitable applications include video on demand and large Web servers.
- *Write-once 3D disks* supporting more than 100 Gbytes per 120-mm disk. Access time to 100-Mbyte blocks of 10–100 milliseconds, with readout rates of more than 500 Mbits/sec. Suitable applications include archiving of data requiring permanent storage yet rapid access, such as medical data and high-resolution maps and satellite images.
- *Pre-recorded 3D disks* supporting more than 100 Gbytes per 120-mm disk and readout rates greater than 200 Mbits/sec. Suitable applications include distributing computer programs, movies, and multimedia. ❖

References

1. D. Psaltis and F. Mok, "Holographic Memories," *Scientific American*, Nov. 1995.
2. J.H. Hong et al., "Volume Holographic Memory Systems: Techniques and Architectures," *Optical Eng.*, Aug. 1995, pp. 2,193-2,203.
3. J.F. Heanue, M.C. Bashaw, and L. Hesselink, "Volume Holographic Storage and Retrieval of Digital Data," *Science*, Aug. 1994, pp. 749-752.
4. Selected papers on Holographic Storage, G.T. Sincerbox, ed., *SPIE Milestone Series*, Vol. MS 95, SPIE, Bellingham, Wash., 1994.
5. A. Pu and D. Psaltis, "High-Density Recording in Photopolymer-Based Holographic Three-Dimensional Disks," *Applied Optics*, May 1996, pp. 2,389-2,398.
6. G. Barbastathis, M. Levene, and D. Psaltis, "Shift Multiplexing with Spherical Reference Waves," *Applied Optics*, May 1996, pp. 2,403-2,417.
7. G.W. Burr, *Volume Holographic Storage Using the 90º Geometry*, doctoral dissertation, California Institute of Technology, Pasadena, Calif., 1996.
8. M.-P. Bernal et al., "A Precision Tester for Studies of Holographic Optical Storage Materials and Recording Physics," *Applied Optics*, May 1996, pp. 2,360-2,374.
9. G.W. Burr et al., "Modulation Coding for Pixel-

Matched Holographic Storage," *Optics Letters*, May 1997, pp. 639-641.
10. F.H. Mok, G.W. Burr, and D. Psaltis, "System Metric for Holographic Memory Systems," *Optics Letters*, June 1996, pp. 896-898.
11. K. Blotekjaer, "Limitations on Holographic Storage Capacity of Photochromic and Photorefractive Media," *Applied Optics*, June 1979, pp. 57-67.
12. M.A. Neifeld and M. McDonald, "Error Correction for Increasing the Usable Capacity of Photorefractive Memories," *Optics Letters*, Sept. 1994, pp. 1,483-1,485.
13. B.H. Olson and S.C. Esener, "Partial Response Precoding for Parallel Readout Optical Memories," *Optics Letters*, May 1994, pp. 661-663.
14. J.-J. P. Drolet et al., "Compact, Integrated Dynamic Holographic Memory with Refreshed Holograms," *Optics Letters*, Apr. 1997, pp. 552-554.

Demetri Psaltis is the Thomas G. Myers Professor of Electrical Engineering and the director of the Neuromorphic Systems Engineering Research Center at the California Institute of Technology. His many research interests include optical information processing and holography. Psaltis received a BSc in electrical engineering and economics and an MSc and a PhD in electrical engineering, all from Carnegie Mellon University. He has authored or co-authored more than 280 publications. He is a fellow of the Optical Society of America and received the International Commission of Optics Prize.

Geoffrey W. Burr researches holographic data storage systems at IBM Almaden Research Center. Other research interests include optical correlation and information processing, optical interconnects, and 3D displays. Burr received a BS from the University of Buffalo, and an MS and a PhD from the California Institute of Technology, all in electrical engineering. He is a member of IEEE, the Optical Society of America, and SPIE.

Contact Psaltis at Dept. of Electrical Eng., MS 136-93, California Inst. of Technology, Pasadena, CA 91125; psaltis@caltech.edu. Contact Burr at IBM Almaden Research Center, D2-K18, 650 Harry Rd., San Jose, CA 95120; burr@almaden.ibm.com.